

Restoring Local Causality and Objective Reality to the Entangled Photons

Joy Christian*

Department of Physics, University of Oxford, Parks Road, Oxford OX1 3PU, United Kingdom

Unlike our basic theories of space and time, quantum mechanics is not a locally causal theory. Moreover, it is widely believed that any hopes of restoring local causality within a realistic theory have been undermined by Bell's theorem and its supporting experiments. By contrast, we provide a strictly local, deterministic, and realistic explanation for the correlations observed in two such supporting experiments, performed independently at Orsay and Innsbruck. To this end, a pair of local variables is constructed to simulate detections of photon polarizations at various angles, chosen freely by Alice and Bob. These generate purely random outcomes, $\mathcal{A} = \pm 1$ and $\mathcal{B} = \pm 1$, occurring within a parallelized 3-sphere. These outcomes do not depend on the contexts of measurements, but are determined entirely by the randomly chosen initial orientation of the 3-sphere. When they are compared, however, the correlation between them turn out to be exactly equal to $-\cos 2(\alpha - \beta)$, with the corresponding CHSH inequality violated for the polarization angles α , α' , β , and β' in precisely the manner predicted by quantum mechanics. The key ingredient in our explanation is the topology of the 3-sphere, which remains closed under multiplication, thus preserving the locality condition of Bell. It allows us to model the physical space as a 3-sphere, and reveals that the illusion of quantum nonlocality in the present case stems from a twist in the Hopf fibration of the 3-sphere.

One of the first steps we often take towards measuring a physical quantity is to set up a Cartesian coordinate system $\{x, y, z\}$ in the Euclidean space \mathbb{E}_3 . This amounts to modeling the Euclidean space as a 3-fold product of the real line, \mathbb{R}^3 . This procedure has become so familiar to us that in practice we often identify \mathbb{E}_3 with its Cartesian model, and simply think of \mathbb{R}^3 as *the* Euclidean space. As we shall see, however, this seemingly innocuous act of convenience comes with a heavy price: It is largely responsible for the illusion of quantum nonlocality [1][2]. In what follows we shall demonstrate that, once a correct coordinate-free geometric model of the physical space is used, the correlations observed in the EPR-inspired experiments involving photon pairs [3][4][5], namely

$$\begin{aligned}\mathcal{A}(\alpha) &= \pm 1, \quad \mathcal{B}(\beta) = \pm 1, \\ \mathcal{E}(\alpha) &= 0, \quad \mathcal{E}(\beta) = 0, \\ \mathcal{E}(\alpha, \beta) &= -\cos 2(\alpha - \beta),\end{aligned}\tag{1}$$

are easily understood, in a strictly local-realistic terms. Here $\mathcal{E}(\alpha, \beta)$ represents the expected value of simultaneously observing remote measurement results $\mathcal{A}(\alpha)$ and $\mathcal{B}(\beta)$ along the polarization angles α and β , respectively.

Euclid himself of course did not think of \mathbb{E}_3 in terms of triples of real numbers. He defined its representation axiomatically, entirely in terms of primitive geometric objects such as points and lines, together with a list of their properties, from which his theorems of geometry follow. Today we know, however, that it is quite problematic to give a suitable definition of Euclidean space in the spirit of Euclid, and hence in physics we instinctively identify \mathbb{E}_3 with \mathbb{R}^3 whenever possible. But there is no natural, geometrically-determined way to identify the two spaces without introducing an *unphysical* notion of arbitrarily distinguished coordinate system [6]. This difficulty is clearly relevant in the study of Bell's theorem [2], for time and again we have learned that surreptitious introduction of unphysical ideas in physics could lead to distorted views of the physical reality. A coordinate-free representation of the Euclidean space is undoubtedly preferable, if what is at stake is the very nature of the physical reality.

Fortunately, precisely such a representation of \mathbb{E}_3 , with a rich algebraic structure, was provided by Grassmann in 1844 [7]. As in Euclid's geometry, the basic elements of this powerful structure are not coordinate systems, but points, lines, planes, and volumes, *all treated on equal footing*. Today one begins this framework by postulating a unit volume element (or a trivector) in \mathbb{E}_3 , defined by

$$I = \mathbf{e}_x \wedge \mathbf{e}_y \wedge \mathbf{e}_z,\tag{2}$$

with $\{\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z\}$ being a set of orthonormal vectors, " \wedge " the outer product of Grassmann, and $I^2 = -1$ [7]. Each vector \mathbf{e}_j is then a solution of the equation $I \wedge \mathbf{e}_j = 0$, and every pair of them respects the fundamental geometric

*Electronic address: joy.christian@wolfson.ox.ac.uk

product

$$\mathbf{e}_j \mathbf{e}_k = \mathbf{e}_j \cdot \mathbf{e}_k + \mathbf{e}_j \wedge \mathbf{e}_k, \quad (3)$$

where “ \cdot ” represents the inner product, and $\mathbf{e}_j \wedge \mathbf{e}_k$ are unit bivectors (which also square to -1) with counterclockwise sense for the cyclicly permuted indices ($j, k = x, y$, or z). The resulting *real* geometric structure is a linear vector space spanned by the basis

$$\{1, \mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z, \mathbf{e}_x \wedge \mathbf{e}_y, \mathbf{e}_y \wedge \mathbf{e}_z, \mathbf{e}_z \wedge \mathbf{e}_x, \mathbf{e}_x \wedge \mathbf{e}_y \wedge \mathbf{e}_z\}, \quad (4)$$

which encodes a graded linear algebra of dimensions eight. This algebra intrinsically characterizes the space \mathbb{E}_3 .

Our interest, however, lies in a certain subalgebra of this algebra, the so-called even subalgebra of dimensions four, defined by the bivector (or spinor) basis [7][6]:

$$\{1, \mathbf{e}_x \wedge \mathbf{e}_y, \mathbf{e}_y \wedge \mathbf{e}_z, \mathbf{e}_z \wedge \mathbf{e}_x\}. \quad (5)$$

Crucially for our purposes, this subalgebra happens to remain closed under multiplication. Consequently, it can be used by itself to model the physical space. In fact, it provides the most natural coordinate-free representation of the physical space as a parallelized 3-sphere [6], which differs from \mathbb{R}^3 only by a single point: $S^3 = \mathbb{R}^3 \cup \{\infty\}$. Both \mathbb{R}^3 and S^3 are three-dimensional manifolds that are parallelized by a vanishing Riemann curvature, but unlike in \mathbb{R}^3 the torsion within S^3 is non-vanishing [8]. S^3 is thus a one-point compactification of \mathbb{R}^3 that remains as flat as \mathbb{R}^3 [9].

The vectors and trivectors are no longer intrinsic to the above subalgebra, but belong to a dual space. Only the scalars and bivectors—*treated on equal footing*—are taken to be intrinsic parts of the subalgebra. This can be seen more clearly if we use the condition $I \wedge \mathbf{e}_j = 0$ to rewrite the basis bivectors defined in Eq. (5) as $I \cdot \mathbf{e}_z$, $I \cdot \mathbf{e}_x$, and $I \cdot \mathbf{e}_y$. Their geometric product, analogous to the one in Eq. (3), then leads to the defining equation of this subalgebra:

$$(I \cdot \mathbf{e}_j)(I \cdot \mathbf{e}_k) = -\delta_{jk} - \epsilon_{jkl} (I \cdot \mathbf{e}_l). \quad (6)$$

Evidently, despite the occurrences of trivectors and basis vectors, only the basis scalar and bivectors are involved in this definition. Consequently, in what follows only scalars and bivectors (and their combinations) will have direct physical significance—vectors and trivectors will merely facilitate computational ease, or “hidden” variables.

Given the bivector basis defined by Eq. (5), any generic bivector such as $I \cdot \mathbf{a}$ can be expanded in this basis as

$$I \cdot \mathbf{a} = \{a_x \mathbf{e}_y \wedge \mathbf{e}_z + a_y \mathbf{e}_z \wedge \mathbf{e}_x + a_z \mathbf{e}_x \wedge \mathbf{e}_y\}. \quad (7)$$

In many ways bivectors—and also vectors and trivectors—behave just like ordinary numbers within this geometrical framework. For this reason Grassmann referred to them as “extensive magnitudes”, and treated them with the same respect as he would treat ordinary numbers. And for the same reason Hestenes today refers to all such quantities as “directed numbers” [7]. They are simply numbers of higher grades and dimensions, with built-in directional attributes.

It is also worth noting that, although there is isomorphism between the vector subspace and the bivector subspace, a bivector is an abstract entity of its own, with properties quite distinct from those of a vector [7]. In fact it is a skeleton of only three properties: (1) a sense, quantified by a $+$ or $-$ sign, indicating whether it represents a counterclockwise or clockwise rotation, (2) a magnitude, which is equal to unity for all bivectors we will be considering, and (3) a direction, which can be specified by a dual vector normal to its shape-independent plane. Thus, despite appearances, neither the trivector I nor the vector \mathbf{a} is an intrinsic part of the bivector $I \cdot \mathbf{a}$. Moreover, it is easy to verify that unit bivectors such as $I \cdot \mathbf{b}$ simply represent intrinsic points of a unit 2-sphere contained within the 3-sphere:

$$\begin{aligned} \|I \cdot \mathbf{b}\|^2 &= (+I \cdot \mathbf{b})(-I \cdot \mathbf{b}) \\ &= -I^2 \mathbf{b} \mathbf{b} = \mathbf{b} \mathbf{b} = \mathbf{b} \cdot \mathbf{b} = \|\mathbf{b}\|^2 = +1. \end{aligned} \quad (8)$$

More precisely, they represent the equatorial points of a parallelized 3-sphere modeling the physical space [8].

Given two such unit bivectors, say $I \cdot \mathbf{a}$ and $I \cdot \mathbf{b}$, the bivector subalgebra (6) leads to the well known identity

$$(I \cdot \mathbf{a})(I \cdot \mathbf{b}) = -\mathbf{a} \cdot \mathbf{b} - I \cdot (\mathbf{a} \times \mathbf{b}), \quad (9)$$

provided we use the duality relation $\mathbf{a} \wedge \mathbf{b} = I \cdot (\mathbf{a} \times \mathbf{b})$ defined between the vector $\mathbf{a} \times \mathbf{b}$ and the bivector $\mathbf{a} \wedge \mathbf{b}$.

The above identity provides a natural representation of points of a parallelized 3-sphere [8]. The bivectors $I \cdot \mathbf{a}$ and $I \cdot \mathbf{b}$ appearing on its LHS represent the equatorial points of the 3-sphere, and the *real* quaternion appearing on its RHS—which is a sum of a scalar and a bivector—represents a non-equatorial point of the same sphere. An equator of a parallelized 3-sphere, however, which is a 2-sphere, does not remain closed under multiplication. The 3-sphere itself, on the other hand, *does* remain closed under multiplication, thereby correctly encoding the topology underlying the subalgebra (6). Conversely, any arbitrary point P of a parallelized 3-sphere can always be factorized into any number of points: $P = ABCD\dots$. Needless to say, this is a highly nontrivial and powerful property of the 3-sphere. To appreciate its non-triviality, consider a product of infinitely many points of a 3-sphere. Such a product will simply be another point of the 3-sphere [8]. By contrast, this will not be true in the case of a 2-sphere even for just two points, as is evident from the above identity. The fact that both 3-sphere and its algebraic representation (9) remain closed under multiplication suggests that we should represent measurement results in the present case by local maps of the form

$$\mathcal{A}(\mathbf{a}, \lambda) : \mathbb{R}^3 \times \Lambda \longrightarrow S^2 \subset S^3, \quad (10)$$

where Λ is a space of complete states λ , vector $\mathbf{a} \in \mathbb{R}^3$ specifies the context of measurement, and S^2 is an equatorial 2-sphere within a unit, parallelized 3-sphere. Then not only the measurement results, but also their products would remain within the 3-sphere, thereby respecting the locality (or factorizability) condition of Bell [2]. We shall see that a local-realistic interplay between the points of such a 3-sphere and its equatorial 2-sphere is what is truly responsible for the EPR correlations manifested in nature.

So far we have considered the bivector subalgebra (6) with arbitrarily fixed basis, as in definition (5). The convention usually is to assume a right-handed set of basis bivectors, and so far we have followed this convention. The algebra itself, however, does not fix the handedness of the basis. We could have equally well started out with a left-handed set of bivectors, by letting $-I$ instead of $+I$ fix the basis. Equation (6) would have then had the alternate form:

$$(-I \cdot \mathbf{e}_j)(-I \cdot \mathbf{e}_k) = -\delta_{jk} - \epsilon_{jkl}(-I \cdot \mathbf{e}_l). \quad (11)$$

Comparing this equation with equation (6) we see that there remains a sign ambiguity in the definition of our subalgebra (*cf.* Refs. [6] and [7]):

$$(I \cdot \mathbf{e}_j)(I \cdot \mathbf{e}_k) = -\delta_{jk} \pm \epsilon_{jkl}(I \cdot \mathbf{e}_l). \quad (12)$$

Consequently, following the time-honored mathematical practice of turning an ambiguity of sign into virtue, we define the handedness of this entire subalgebra as our “hidden variable.” In other words, we specify the complete state of the photons we are about to study as $\boldsymbol{\mu} = \pm I$, thereby defining the basis of our entire subalgebra by the equation

$$(\boldsymbol{\mu} \cdot \mathbf{e}_j)(\boldsymbol{\mu} \cdot \mathbf{e}_k) = -\delta_{jk} - \epsilon_{jkl}(\boldsymbol{\mu} \cdot \mathbf{e}_l). \quad (13)$$

The identity (9) for the generic bivectors then becomes

$$(\boldsymbol{\mu} \cdot \mathbf{a})(\boldsymbol{\mu} \cdot \mathbf{b}) = -\mathbf{a} \cdot \mathbf{b} - \boldsymbol{\mu} \cdot (\mathbf{a} \times \mathbf{b}), \quad (14)$$

along with *indefinite* duality relation $\mathbf{a} \wedge \mathbf{b} = \boldsymbol{\mu} \cdot (\mathbf{a} \times \mathbf{b})$. In other words, the duality between the wedge product and cross product within our subalgebra alternates between the right-hand and left-hand rules [8]. Clearly, then, our complete state $\boldsymbol{\mu} = \pm I$ represents a far deeper hidden structure than the variables considered by Bell [2]. It un-fixes the orientation of the entire physical space S^3 , and turns it into a shared randomness between Alice and Bob.

We are now well equipped to take up the question of EPR correlations exhibited by the pair of entangled photons in the actual experiments [4][5]. This question, of course, has been well scrutinized in the literature [3]. We will restrict to the most basic aspects of the question, and follow its treatment given in Refs. [3] and [5]. In a quantum mechanical description of the experiment involving photon pairs one usually assumes that the system has been prepared in the singlet state

$$|\Psi_{-}\rangle = \frac{1}{\sqrt{2}} \left\{ |H\rangle_1 \otimes |V\rangle_2 - |V\rangle_1 \otimes |H\rangle_2 \right\}, \quad (15)$$

where $|H\rangle$ and $|V\rangle$ denote the horizontal and vertical polarization states of the photons along the directions \mathbf{e}_x and \mathbf{e}_y , respectively; and the subscripts 1 and 2 refer to the photons 1 and 2, respectively. The photons are thus assumed

to be propagating in the \mathbf{e}_z direction. We could equally well consider the polarization state $|\Psi_+\rangle$, but that would not add anything significant to our concerns here. Both polarization states, $|\Psi_+\rangle$ and $|\Psi_-\rangle$, are invariant under rotations about the \mathbf{e}_z axis, but the state $|\Psi_+\rangle$ is even under reflections, whereas the state $|\Psi_-\rangle$ is odd under reflections.

In a typical experimental run Alice and Bob measure polarizations of the photons along two different directions in the plane perpendicular to the \mathbf{e}_z axis. Alice measures polarizations along the direction \mathbf{a} , which makes an angle α with the \mathbf{e}_x axis, whereas Bob measures polarizations along the direction \mathbf{b} , which makes an angle β with the \mathbf{e}_x axis. Individually, the binary results observed by Alice and Bob, namely $A(\alpha) = \pm 1$ and $B(\beta) = \pm 1$, are found to be completely random, with equal probabilities for the outcomes $+1$ and -1 . When these results are compared, however, they are found to be strongly correlated, in agreement with the quantum mechanical predictions we have summarized in Eq. (1).

Our goal now is to reproduce these quantum mechanical predictions *exactly*, within the geometrical model of the physical space discussed above. To this end, we have assumed that the complete state of the photons is given by $\boldsymbol{\mu} = \pm I$, where I is the fundamental trivector defined in Eq. (2). The detections of photon polarizations observed by Alice and Bob along their respective axes \mathbf{a} and \mathbf{b} , with the bivector basis fixed by the trivector $\boldsymbol{\mu}$, can then be represented *intrinsically* as points of the physical space S^3 , by the following two local variables:

$$S^3 \ni \mathcal{A}(\alpha, \boldsymbol{\mu}) = (-I \cdot \tilde{\mathbf{a}})(+\boldsymbol{\mu} \cdot \tilde{\mathbf{a}}) = \begin{cases} +1 & \text{if } \boldsymbol{\mu} = +I \\ -1 & \text{if } \boldsymbol{\mu} = -I \end{cases} \quad (16)$$

and

$$S^3 \ni \mathcal{B}(\beta, \boldsymbol{\mu}) = (+I \cdot \tilde{\mathbf{b}})(+\boldsymbol{\mu} \cdot \tilde{\mathbf{b}}) = \begin{cases} -1 & \text{if } \boldsymbol{\mu} = +I \\ +1 & \text{if } \boldsymbol{\mu} = -I \end{cases}, \quad (17)$$

with equal probabilities for $\boldsymbol{\mu}$ being either $+I$ or $-I$, and the rotating vectors $\tilde{\mathbf{a}}$ and $\tilde{\mathbf{b}}$ defined as

$$\begin{aligned} \tilde{\mathbf{a}} &= \mathbf{e}_x \cos 2\alpha + \mathbf{e}_y \sin 2\alpha, \\ \text{and } \tilde{\mathbf{b}} &= \mathbf{e}_x \cos 2\beta + \mathbf{e}_y \sin 2\beta. \end{aligned} \quad (18)$$

Note that $\mathcal{A}(\alpha, \boldsymbol{\mu})$ and $\mathcal{B}(\beta, \boldsymbol{\mu})$, in addition to being manifestly realistic, are strictly *local* variables¹. In fact, they are not even contextual [10]. Alice's measurement result, although refers to a freely chosen angle α , depends only on the initial state $\boldsymbol{\mu}$; and likewise, Bob's measurement result, although refers to a freely chosen angle β , depends only on the initial state $\boldsymbol{\mu}$. In other words, all possible measurement results at all possible angles are completely determined by the initial orientation of the physical space specified by $\boldsymbol{\mu}$, or equivalently by that of the 3-sphere. Moreover, it is easy to check using the identity (14) that $\mathcal{A}(\alpha, \boldsymbol{\mu})$ and $\mathcal{B}(\beta, \boldsymbol{\mu})$ are purely binary numbers. For consider the product

$$\begin{aligned} (+I \cdot \tilde{\mathbf{a}})(+I \cdot \tilde{\mathbf{a}}') &= -\tilde{\mathbf{a}} \cdot \tilde{\mathbf{a}}' - I \cdot (\tilde{\mathbf{a}} \times \tilde{\mathbf{a}}') \\ &= -\cos \theta_{\tilde{\mathbf{a}}\tilde{\mathbf{a}}'} - (I \cdot \tilde{\mathbf{c}}) \sin \theta_{\tilde{\mathbf{a}}\tilde{\mathbf{a}}'}, \end{aligned} \quad (19)$$

where $\tilde{\mathbf{c}} = \tilde{\mathbf{a}} \times \tilde{\mathbf{a}}' / |\tilde{\mathbf{a}} \times \tilde{\mathbf{a}}'|$, and $\theta_{\tilde{\mathbf{a}}\tilde{\mathbf{a}}'}$ is the angle between $\tilde{\mathbf{a}}$ and $\tilde{\mathbf{a}}'$. The RHS of this identity is a quaternion, which represents a non-equatorial point of the parallelized 3-sphere [8]. Physically, it represents a rotation by angle $2\theta_{\tilde{\mathbf{a}}\tilde{\mathbf{a}}'}$ about the $\tilde{\mathbf{c}}$ -axis. The limit $\tilde{\mathbf{a}}' \rightarrow \tilde{\mathbf{a}}$ then gives $(+I \cdot \tilde{\mathbf{a}})(+I \cdot \tilde{\mathbf{a}}) = -1$, which shows that unit bivectors square to -1 . Moreover, in equation (8) we have already seen the conjugate limiting case; namely, $(-I \cdot \tilde{\mathbf{a}})(+I \cdot \tilde{\mathbf{a}}) = +1$. These two cases show that the pairs of numbers $\{+1, -1\}$ are intrinsic parts of the parallelized 3-sphere. They are in fact antipodal points of one of its equators, representing the set of measurement results observed by our experimenters.

In statistical terms, however, the above variables are raw scores as opposed to standard scores [11]. Recall that a standard score indicates how many standard deviations the observation or datum is above or below the mean. If x is a raw (or unnormalized) score and \bar{x} is its mean value, then the standard (or normalized) score, $z(x)$, is defined by

$$z(x) = \frac{x - \bar{x}}{\sigma(x)}, \quad (20)$$

¹ Needless to say, $\mathcal{A}(\alpha, \boldsymbol{\mu})$ and $\mathcal{B}(\beta, \boldsymbol{\mu})$ are two *different* functions of the random variable $\boldsymbol{\mu}$. Moreover, they are *statistically independent events* occurring within a 3-sphere, with factorized joint probability $P(\mathcal{A} \text{ and } \mathcal{B}) = P(\mathcal{A}) \times P(\mathcal{B}) \leq \frac{1}{2}$. Therefore their product $\mathcal{A}\mathcal{B}$ is guaranteed to be equal to -1 only for the special case $\alpha = \beta$. For all other α and β , $\mathcal{A}\mathcal{B}$ will alternate between the values -1 and $+1$.

where $\sigma(x)$ is the standard deviation of x . A standard score thus represents the distance between a raw score and the population mean in the units of standard deviation, and allows one to make comparisons of raw scores that come from very different sources. In other words, the mean value of the standard score itself is always zero, with standard deviation unity. In terms of these concepts the bivariate correlation between raw scores x and y is defined as

$$\mathcal{E}(x, y) = \frac{\lim_{n \gg 1} \left[\frac{1}{n} \sum_{i=1}^n (x^i - \bar{x}) (y^i - \bar{y}) \right]}{\sigma(x) \sigma(y)} \quad (21)$$

$$= \lim_{n \gg 1} \left[\frac{1}{n} \sum_{i=1}^n z(x^i) z(y^i) \right]. \quad (22)$$

It is vital to appreciate that covariance by itself—*i.e.*, the numerator of the equation (21) by itself—does not provide the correct measure of association between the raw scores, not the least because it depends on different units and scales (or different scales of dispersion) that may have been used (advertently or inadvertently) in the measurements of such scores [11]. Therefore, to arrive at the correct measure of association between the raw scores one must either use equation (21), with the product of standard deviations in the denominator, or use covariance of the standardized variables, as in Eq. (22).

These basic statistical concepts are crucial for understanding our local-realistic explanation of the EPR correlations. As we saw above, the variables $\mathcal{A}(\alpha, \boldsymbol{\mu})$ and $\mathcal{B}(\beta, \boldsymbol{\mu})$ are pure binary numbers, albeit occurring as antipodal points within a parallelized 3-sphere. As random variables, however, they are products of two factors—one random and another non-random. For instance, within $\mathcal{A}(\alpha, \boldsymbol{\mu})$ the factor $(+\boldsymbol{\mu} \cdot \tilde{\mathbf{a}})$ is a random bivector—a function of the hidden variable $\boldsymbol{\mu}$, whereas the factor $(-I \cdot \tilde{\mathbf{a}})$ is a non-random bivector, independent of the hidden variable $\boldsymbol{\mu}$. Consequently, as a function, each of the numbers $\mathcal{A}(\alpha, \boldsymbol{\mu})$ and $\mathcal{B}(\beta, \boldsymbol{\mu})$ is generated with a *different* standard deviation—*i.e.*, a *different* size of a typical error. More specifically, the number $\mathcal{A}(\alpha, \boldsymbol{\mu})$ is generated with the standard deviation $(-I \cdot \tilde{\mathbf{a}})$, whereas the number $\mathcal{B}(\beta, \boldsymbol{\mu})$ is generated with the standard deviation $(+I \cdot \tilde{\mathbf{b}})$. These deviations can be calculated easily. Since errors in a linear relation such as (16) propagate linearly, the standard deviation of $\mathcal{A}(\alpha, \boldsymbol{\mu})$ is equal to $(-I \cdot \tilde{\mathbf{a}})$ times the standard deviation of $(+\boldsymbol{\mu} \cdot \tilde{\mathbf{a}})$ (which we write as $\sigma(A)$), and similarly the standard deviation of $\mathcal{B}(\beta, \boldsymbol{\mu})$ is equal to $(+I \cdot \tilde{\mathbf{b}})$ times the standard deviation of $(+\boldsymbol{\mu} \cdot \tilde{\mathbf{b}})$ (which we write as $\sigma(B)$):

$$\begin{aligned} \sigma(\mathcal{A}) &= (-I \cdot \tilde{\mathbf{a}}) \sigma(A) \\ \text{and } \sigma(\mathcal{B}) &= (+I \cdot \tilde{\mathbf{b}}) \sigma(B). \end{aligned} \quad (23)$$

But since all bivectors we have been considering are normalized to unity, and since the mean of $(+\boldsymbol{\mu} \cdot \tilde{\mathbf{a}})$ vanishes on the account of $\boldsymbol{\mu}$ being a fair coin, its standard deviation is easy to calculate, and it turns out to be equal to unity:

$$\begin{aligned} \sigma(A) &= \sqrt{\frac{1}{n} \sum_{i=1}^n \left\| A(\alpha, \boldsymbol{\mu}^i) - \overline{A(\alpha, \boldsymbol{\mu}^i)} \right\|^2} \\ &= \sqrt{\frac{1}{n} \sum_{i=1}^n \left\| (+\boldsymbol{\mu}^i \cdot \tilde{\mathbf{a}}) - 0 \right\|^2} = 1, \end{aligned} \quad (24)$$

where the last equality follows from the fact that all bivectors $(+\boldsymbol{\mu}^i \cdot \tilde{\mathbf{a}})$ are normalized to unity. Similarly, we find that the standard deviation $\sigma(B)$ is also equal to 1. As a result, the standard deviation of $\mathcal{A}(\alpha, \boldsymbol{\mu})$ works out to be equal to $(-I \cdot \tilde{\mathbf{a}})$, and the standard deviation of $\mathcal{B}(\beta, \boldsymbol{\mu})$ works out to be equal to $(+I \cdot \tilde{\mathbf{b}})$. Putting these two results together, we arrive at the following standard scores corresponding to the raw scores (16) and (17):

$$\begin{aligned} A(\alpha, \boldsymbol{\mu}) &= \frac{\mathcal{A}(\alpha, \boldsymbol{\mu}) - \overline{\mathcal{A}(\alpha, \boldsymbol{\mu})}}{\sigma(\mathcal{A})} \\ &= \frac{\mathcal{A}(\alpha, \boldsymbol{\mu}) - 0}{(-I \cdot \tilde{\mathbf{a}})} = (+\boldsymbol{\mu} \cdot \tilde{\mathbf{a}}) \end{aligned} \quad (25)$$

and

$$\begin{aligned} B(\beta, \boldsymbol{\mu}) &= \frac{\mathcal{B}(\beta, \boldsymbol{\mu}) - \overline{\mathcal{B}(\beta, \boldsymbol{\mu})}}{\sigma(\mathcal{B})} \\ &= \frac{\mathcal{B}(\beta, \boldsymbol{\mu}) - 0}{(+I \cdot \tilde{\mathbf{b}})} = (+\boldsymbol{\mu} \cdot \tilde{\mathbf{b}}), \end{aligned} \quad (26)$$

where we have used the identities $(+I \cdot \tilde{\mathbf{a}})(-I \cdot \tilde{\mathbf{a}}) = +1$ and $(-I \cdot \tilde{\mathbf{b}})(+I \cdot \tilde{\mathbf{b}}) = +1$, respectively (see Appendix for the formal basis of these results).

Not surprisingly, just like the raw scores $\mathcal{A}(\alpha, \boldsymbol{\mu})$ and $\mathcal{B}(\beta, \boldsymbol{\mu})$, these standard scores are also strictly *local* variables: $A(\alpha, \boldsymbol{\mu})$ depends only on the freely chosen angle α and the initial state $\boldsymbol{\mu}$, and likewise $B(\beta, \boldsymbol{\mu})$ depends only on the freely chosen angle β and the initial state $\boldsymbol{\mu}$. Moreover, despite appearances, $A(\alpha, \boldsymbol{\mu})$ and $B(\beta, \boldsymbol{\mu})$ are simply binary measurement results, ± 1 , albeit occurring within the compact topology of the 3-sphere rather than the real line:

$$S^3 \supset S^2 \ni A(\alpha, \boldsymbol{\mu}) = (+\boldsymbol{\mu} \cdot \tilde{\mathbf{a}}) = \pm 1 \text{ about } \tilde{\mathbf{a}} \in \mathbb{R}^3, \quad (27)$$

$$S^3 \supset S^2 \ni B(\beta, \boldsymbol{\mu}) = (+\boldsymbol{\mu} \cdot \tilde{\mathbf{b}}) = \pm 1 \text{ about } \tilde{\mathbf{b}} \in \mathbb{R}^3. \quad (28)$$

In fact, since the space of all bivectors $(+\boldsymbol{\mu} \cdot \tilde{\mathbf{a}})$ is isomorphic to the equatorial 2-sphere contained within the 3-sphere [12], each standard score $A(\alpha, \boldsymbol{\mu})$ of Alice—corresponding to a predetermined element of reality of the photon—is uniquely identified with a definite point of this 2-sphere. In other words, every predetermined polarization of a given photon gives rise to a unique standard score $A(\alpha, \boldsymbol{\mu})$, which in turn is unambiguously represented by a definite point of the equatorial 2-sphere, and likewise for the standard scores $B(\beta, \boldsymbol{\mu})$ of Bob [8].

Now, since we have assumed that initially there was 50/50 chance between the right-handed and left-handed orientations of the physical space—*i.e.*, equal chance between the initial states $\boldsymbol{\mu} = +I$ and $\boldsymbol{\mu} = -I$, the expectation values of the local outcomes trivially work out to be

$$\mathcal{E}(\theta) = \lim_{n \gg 1} \left[\frac{1}{n} \sum_{i=1}^n \mathcal{A}(\theta, \boldsymbol{\mu}^i) \right] = \overline{\mathcal{A}(\theta, \boldsymbol{\mu})} = 0, \quad (29)$$

where $\theta = \alpha$ or β . On the other hand, to determine the *correct* correlation between the joint observations of Alice and Bob we must calculate covariance between the standard scores $A(\alpha, \boldsymbol{\mu})$ and $B(\beta, \boldsymbol{\mu})$, not the raw scores:

$$\mathcal{E}(\alpha, \beta) = \lim_{n \gg 1} \left[\frac{1}{n} \sum_{i=1}^n A(\alpha, \boldsymbol{\mu}^i) B(\beta, \boldsymbol{\mu}^i) \right]. \quad (30)$$

Conversely, the covariance of the raw scores will not give us the correct correlation between the observations of Alice and Bob, because $\mathcal{A}(\alpha, \boldsymbol{\mu})$ and $\mathcal{B}(\beta, \boldsymbol{\mu})$ have been generated with *different* scales of dispersion for each direction, and hence they are subject to *different* standard deviations or random errors for each direction [11]. On the other hand, the standard scores—by definition—have been standardized for all kinds of possible random errors. Moreover, their product $A(\alpha, \boldsymbol{\mu}) B(\beta, \boldsymbol{\mu})$ —as a non-equatorial point of the 3-sphere—represents a definite binary value:

$$S^3 \ni A(\alpha, \boldsymbol{\mu}) B(\beta, \boldsymbol{\mu}) = \pm 1 \text{ about a direction in } \mathbb{R}^4. \quad (31)$$

In fact, our 3-sphere is entirely made of such product points, each of definite value $+1$ or -1 , depending on its orientation defined by $\boldsymbol{\mu}$. More precisely, the space of all such product points $P = AB$ is homotopic to a round 3-sphere: $P_1^2 + P_2^2 + P_3^2 + P_4^2 = 1$ [8][12]. This can be seen more clearly if we expand AB using the identity (14):

$$\begin{aligned} A(\alpha, \boldsymbol{\mu}) B(\beta, \boldsymbol{\mu}) &= (+\boldsymbol{\mu} \cdot \tilde{\mathbf{a}})(+\boldsymbol{\mu} \cdot \tilde{\mathbf{b}}) \\ &= -\tilde{\mathbf{a}} \cdot \tilde{\mathbf{b}} - \boldsymbol{\mu} \cdot (\tilde{\mathbf{a}} \times \tilde{\mathbf{b}}) \\ &= -\cos 2(\alpha - \beta) + (\boldsymbol{\mu} \cdot \mathbf{e}_z) \sin 2(\alpha - \beta). \end{aligned} \quad (32)$$

Evidently, the product AB describes a circle of points within S^3 [12], each of definite value $+1$ or -1 , depending on α , β , and $\boldsymbol{\mu} \cdot \mathbf{e}_z \in S^2 \subset S^3$. Substituting this parameterization of AB into Eq. (30), the correlation between the raw

scores $\mathcal{A}(\alpha, \boldsymbol{\mu})$ and $\mathcal{B}(\beta, \boldsymbol{\mu})$ can now be easily seen to reproduce the corresponding quantum mechanical prediction:

$$\begin{aligned}\mathcal{E}(\alpha, \beta) &= -\cos 2(\alpha - \beta) + \lim_{n \gg 1} \left[\frac{1}{n} \sum_{i=1}^n (\boldsymbol{\mu}^i \cdot \mathbf{e}_z) \sin 2(\alpha - \beta) \right] \\ &= -\cos 2(\alpha - \beta) + 0.\end{aligned}\quad (33)$$

Here the summation over microstates $\boldsymbol{\mu}^i$ has only formal significance since operationally \mathbf{e}_z is a “third” direction, exclusive to both \mathbf{a} and \mathbf{b} . If measurements along \mathbf{a} and \mathbf{b} have yielded non-null polarizations, then simultaneous measurement along \mathbf{e}_z could only yield a null result.

Suppose now we consider four possible polarization directions, \mathbf{a} , \mathbf{a}' , \mathbf{b} , and \mathbf{b}' . Then, as is well known, the corresponding CHSH string of expectation values [13], namely

$$\mathcal{E}(\mathbf{a}, \mathbf{b}) + \mathcal{E}(\mathbf{a}, \mathbf{b}') + \mathcal{E}(\mathbf{a}', \mathbf{b}) - \mathcal{E}(\mathbf{a}', \mathbf{b}'), \quad (34)$$

can be rewritten in terms of the products of our local functions $A(\mathbf{a}, \boldsymbol{\mu})$ and $B(\mathbf{b}, \boldsymbol{\mu})$ as

$$\lim_{n \gg 1} \left[\frac{1}{n} \sum_{i=1}^n \{A_{\mathbf{a}}(\boldsymbol{\mu}^i) B_{\mathbf{b}}(\boldsymbol{\mu}^i) + A_{\mathbf{a}}(\boldsymbol{\mu}^i) B_{\mathbf{b}'}(\boldsymbol{\mu}^i) + A_{\mathbf{a}'}(\boldsymbol{\mu}^i) B_{\mathbf{b}}(\boldsymbol{\mu}^i) - A_{\mathbf{a}'}(\boldsymbol{\mu}^i) B_{\mathbf{b}'}(\boldsymbol{\mu}^i)\} \right]. \quad (35)$$

But since the standard scores $A_{\mathbf{a}}(\boldsymbol{\mu})$ and $B_{\mathbf{b}}(\boldsymbol{\mu})$ represent two independent points of S^3 , we can take them to belong to two disconnected “sections” of S^3 (*i.e.*, two disconnected 2-spheres within S^3), satisfying

$$[A_{\mathbf{n}}(\boldsymbol{\mu}), B_{\mathbf{n}'}(\boldsymbol{\mu})] = 0 \quad \forall \mathbf{n} \text{ and } \mathbf{n}' \in \mathbb{R}^3, \quad (36)$$

which is operationally equivalent to assuming a null result along the direction $\mathbf{n} \times \mathbf{n}'$ exclusive to both \mathbf{n} and \mathbf{n}' . If we now square the integrand of Eq. (35), use the above commutation relations, and use the fact that, by definition, all local functions square to unity (the algebra goes through even when the squares of the local functions are allowed to be -1), then the absolute value of the CHSH string leads to the following variance inequality:

$$\begin{aligned}& |\mathcal{E}(\mathbf{a}, \mathbf{b}) + \mathcal{E}(\mathbf{a}, \mathbf{b}') + \mathcal{E}(\mathbf{a}', \mathbf{b}) - \mathcal{E}(\mathbf{a}', \mathbf{b}')| \\ & \leq \sqrt{\lim_{n \gg 1} \left[\frac{1}{n} \sum_{i=1}^n \{4 + [A_{\mathbf{a}}(\boldsymbol{\mu}^i), A_{\mathbf{a}'}(\boldsymbol{\mu}^i)][B_{\mathbf{b}'}(\boldsymbol{\mu}^i), B_{\mathbf{b}}(\boldsymbol{\mu}^i)]\} \right]}.\end{aligned}\quad (37)$$

Here the classical commutators $[A_{\mathbf{a}}(\boldsymbol{\mu}^i), A_{\mathbf{a}'}(\boldsymbol{\mu}^i)]$ and $[B_{\mathbf{b}'}(\boldsymbol{\mu}^i), B_{\mathbf{b}}(\boldsymbol{\mu}^i)]$ are simply geometric measures of the parallelizing torsion within the 3-sphere [8]. Next, using the definitions (25) and (26) for the variables $A_{\mathbf{a}}(\boldsymbol{\mu})$ and $B_{\mathbf{b}}(\boldsymbol{\mu})$ and making a repeated use of the identity $(+\boldsymbol{\mu} \cdot \tilde{\mathbf{a}})(+\boldsymbol{\mu} \cdot \tilde{\mathbf{a}}') = -\tilde{\mathbf{a}} \cdot \tilde{\mathbf{a}}' - \boldsymbol{\mu} \cdot (\tilde{\mathbf{a}} \times \tilde{\mathbf{a}}')$, the above inequality can be simplified to

$$\begin{aligned}& |\mathcal{E}(\mathbf{a}, \mathbf{b}) + \mathcal{E}(\mathbf{a}, \mathbf{b}') + \mathcal{E}(\mathbf{a}', \mathbf{b}) - \mathcal{E}(\mathbf{a}', \mathbf{b}')| \\ & \leq \sqrt{\lim_{n \gg 1} \left[\frac{1}{n} \sum_{i=1}^n \{4 + 4[-(\tilde{\mathbf{a}} \times \tilde{\mathbf{a}}') \cdot (\tilde{\mathbf{b}}' \times \tilde{\mathbf{b}}) - \boldsymbol{\mu}^i \cdot \tilde{\mathbf{0}}]\} \right]} \\ & \leq 2 \sqrt{1 - (\tilde{\mathbf{a}} \times \tilde{\mathbf{a}}') \cdot (\tilde{\mathbf{b}}' \times \tilde{\mathbf{b}})},\end{aligned}\quad (38)$$

where $\boldsymbol{\mu}^i \cdot \tilde{\mathbf{0}}$ is a null bivector. It is worth noting here that—because of the nullity of $\boldsymbol{\mu}^i \cdot \tilde{\mathbf{0}}$ —the summation over the microstates $\boldsymbol{\mu}^i$ is not needed to arrive at the last inequality—the summation simply becomes redundant, bringing out the purely geometrical character of the EPR correlations [8] in addition to conforming the consistency of assumption (36). Moreover, using the definitions (18) for the vectors $\tilde{\mathbf{a}}$ and $\tilde{\mathbf{b}}$, the last inequality can be further simplified to

$$\begin{aligned}& |\mathcal{E}(\alpha, \beta) + \mathcal{E}(\alpha, \beta') + \mathcal{E}(\alpha', \beta) - \mathcal{E}(\alpha', \beta')| \\ & \leq 2 \sqrt{1 + \sin 2(\alpha - \alpha') \sin 2(\beta - \beta')}.\end{aligned}\quad (39)$$

It is now easy to see that a maximum violation of $2\sqrt{2}$ of the Bell-CHSH inequality can be achieved for the angles $(\alpha, \alpha', \beta, \beta') = (0^\circ, 45^\circ, 22.5^\circ, 67.5^\circ)$, precisely as observed in the experiments [4][5]. On the other hand, by using

$$-1 \leq \sin 2(\alpha - \alpha') \sin 2(\beta - \beta') \leq +1, \quad (40)$$

the above inequality can be reduced to the form

$$|\mathcal{E}(\alpha, \beta) + \mathcal{E}(\alpha, \beta') + \mathcal{E}(\alpha', \beta) - \mathcal{E}(\alpha', \beta')| \leq 2\sqrt{2}, \quad (41)$$

which is precisely what is predicted by quantum mechanics, and observed in the experiments [4][5].

We are now in a position to explain where exactly the illusion of quantum nonlocality stems from [2]. The key to our explanation is the topology of the parallelized 3-sphere, or equivalently the algebra of rotations in the physical space [7][8]. As we saw in equation (32), topologically 3-sphere is simply a 2-sphere worth of circles, but with a “twist” in the bundle [14][15][16]. That is to say, as an S^1 fiber bundle over S^2 , $S^3 \neq S^2 \times S^1$. And this twist—analogue to the one in a Möbius strip—is responsible for producing the right combination of polarizations, namely $++$, $--$, $+-$, and $-+$, observed in the experiments [4][5]. To bring this out explicitly, we have modeled the physical space, not as \mathbb{R}^3 , but S^3 , in terms of coordinate-free geometry of points and planes introduced by Grassmann [7]. This, in turn, has allowed us to introduce the raw scores observed by Alice and Bob as $\mathcal{A}(\alpha, \boldsymbol{\mu}) = (-I \cdot \tilde{\mathbf{a}})(+\boldsymbol{\mu} \cdot \tilde{\mathbf{a}}) = \pm 1$ and $\mathcal{B}(\beta, \boldsymbol{\mu}) = (+I \cdot \tilde{\mathbf{b}})(+\boldsymbol{\mu} \cdot \tilde{\mathbf{b}}) = \pm 1$, respectively, with the random factors $(+\boldsymbol{\mu} \cdot \tilde{\mathbf{a}})$ and $(+\boldsymbol{\mu} \cdot \tilde{\mathbf{b}})$ representing the photon polarizations, and the fixed factors $(-I \cdot \tilde{\mathbf{a}})$ and $(+I \cdot \tilde{\mathbf{b}})$ representing the analyzers used by Alice and Bob to measure these polarizations. Furthermore, we have taken the randomness $\boldsymbol{\mu} = +I$ or $-I$ shared by Alice and Bob to be the initial orientation (or handedness) of the entire physical space, or equivalently that of a 3-sphere. Consequently, once $\boldsymbol{\mu}$ is given as an initial state, the polarizations along all directions chosen by Alice and Bob would have the same value, because $\boldsymbol{\mu}$ completely fixes the sense of bivectors $\boldsymbol{\mu} \cdot \mathbf{n}$ belonging to $S^2 \subset S^3$, regardless of direction. However, and this is an important point, the polarization $(+\boldsymbol{\mu} \cdot \tilde{\mathbf{a}})$ observed by Alice is measured with respect to the analyzer $(-I \cdot \tilde{\mathbf{a}})$, whereas the polarization $(+\boldsymbol{\mu} \cdot \tilde{\mathbf{b}})$ observed by Bob is measured with respect to the analyzer $(+I \cdot \tilde{\mathbf{b}})$.

To appreciate the significance of the last point, let us begin with the scenario in which Alice and Bob are oblivious of each other’s existence. Then, as is evident from equations (16) and (17), their measurement results would depend only on the sense of $\boldsymbol{\mu}$. Consequently, Alice, for example, would be justified in concluding from $\mathcal{A}(\alpha, \boldsymbol{\mu}) = +1$ that the local handedness associated with the points of S^3 is of counterclockwise variety, since $(-I \cdot \tilde{\mathbf{a}})(+\boldsymbol{\mu} \cdot \tilde{\mathbf{a}}) = +1$ holds for any direction $\tilde{\mathbf{a}}$ chosen by Alice if $\boldsymbol{\mu} = +I$; and, $\boldsymbol{\mu}$, as we noted, defines a consistent sense of handedness over the whole of S^3 for each run of the experiment. Similarly, Bob too would be justified in concluding from $\mathcal{B}(\beta, \boldsymbol{\mu}) = -1$ that the points of the 3-sphere are oriented in the counterclockwise sense, since $(+I \cdot \tilde{\mathbf{b}})(+\boldsymbol{\mu} \cdot \tilde{\mathbf{b}}) = -1$ holds for any direction $\tilde{\mathbf{b}}$ chosen by Bob if $\boldsymbol{\mu} = +I$. Furthermore, these conclusions are unproblematic in the important case of “perfect anti-correlation”—i.e., when $\tilde{\mathbf{b}}$ happens to be equal to $\tilde{\mathbf{a}}$. But since we have $\mathcal{A} = +1$ and $\mathcal{B} = -1$ in this case, this is, in fact, the case “ $+-$ ” observed in the experiments. Analogously, for the initial orientation $\boldsymbol{\mu} = -I$ the results of Alice and Bob would be $\mathcal{A} = -1$ and $\mathcal{B} = +1$, and that would be the case “ $-+$ ” observed in the experiments (with equal probability). Note, however, that so far we have no indication of “nonlocality.”

But now suppose we let $\tilde{\mathbf{b}} \rightarrow -\tilde{\mathbf{a}}$. Then $(+I \cdot \tilde{\mathbf{b}})$ becomes $\{+I \cdot (-\tilde{\mathbf{a}})\} = (-I \cdot \tilde{\mathbf{a}})$, and—unlike in the previous case—the analyzers used by Alice and Bob become mathematically identical, with the corresponding raw scores \mathcal{A} and \mathcal{B} generated with the same scale of dispersion for each direction. As a result, now it would be impossible for Alice and Bob to observe opposite polarizations without violating the consistency of handedness defined by $\boldsymbol{\mu}$ over the whole of S^3 . Thus, if Alice’s result turns out to be $(-I \cdot \tilde{\mathbf{a}})(+\boldsymbol{\mu} \cdot \tilde{\mathbf{a}})$ as before, then Bob’s result must also be $(-I \cdot \tilde{\mathbf{a}})(+\boldsymbol{\mu} \cdot \tilde{\mathbf{a}})$, albeit for the detection of polarization along the angle $\alpha + \frac{\pi}{2}$. In other words, the consistency of handedness throughout the 3-sphere necessitates that the only possible results Alice and Bob could observe in this case are either $\mathcal{A} = +1$ and $\mathcal{B} = +1$, or $\mathcal{A} = -1$ and $\mathcal{B} = -1$, depending on whether the initial orientation for the run has been $\boldsymbol{\mu} = +I$ or $-I$. These are, then, the cases “ $++$ ” and “ $--$ ” observed in the experiments (again, with equal probabilities).

Following Bell [2], such instantaneous changes in the relative measurement outcomes observed at remote stations are usually taken as evidence of “quantum nonlocality” [17]. In our spherical model of the physical space, however, such changes are determined entirely by the intrinsic geometrical and topological structures of the 3-sphere (or equivalently by the algebra of rotations in the physical space) without requiring any form of communication between Alice and Bob. To appreciate this in detail, let us express the results of Alice and Bob as limiting cases of two quaternions

constituting the 3-sphere:

$$S^3 \ni \mathcal{A}(\alpha, \boldsymbol{\mu}) = \lim_{\tilde{\mathbf{a}}' \rightarrow \tilde{\mathbf{a}}} [(-I \cdot \tilde{\mathbf{a}})(+\boldsymbol{\mu} \cdot \tilde{\mathbf{a}}')] = +\lambda \quad (42)$$

and

$$S^3 \ni \mathcal{B}(\beta, \boldsymbol{\mu}) = \lim_{\tilde{\mathbf{b}}' \rightarrow \tilde{\mathbf{b}}} [(+I \cdot \tilde{\mathbf{b}})(+\boldsymbol{\mu} \cdot \tilde{\mathbf{b}}')] = -\lambda, \quad (43)$$

where $\lambda = \pm 1$, with $\boldsymbol{\mu} = \lambda I$, and we assume that the angles $\theta_{\tilde{\mathbf{a}}\tilde{\mathbf{a}}'}$ and $\theta_{\tilde{\mathbf{b}}\tilde{\mathbf{b}}'}$, between $\tilde{\mathbf{a}}$ and $\tilde{\mathbf{a}}'$ and $\tilde{\mathbf{b}}$ and $\tilde{\mathbf{b}}'$, respectively, are infinitesimally small. Moreover, as we saw in Eq. (19), the quaternion $(+I \cdot \tilde{\mathbf{a}})(+\boldsymbol{\mu} \cdot \tilde{\mathbf{a}}')$ can be decomposed as

$$\begin{aligned} (+I \cdot \tilde{\mathbf{a}})(+\boldsymbol{\mu} \cdot \tilde{\mathbf{a}}') &= -\lambda \{ \tilde{\mathbf{a}} \cdot \tilde{\mathbf{a}}' + I \cdot (\tilde{\mathbf{a}} \times \tilde{\mathbf{a}}') \} \\ &= -\lambda \{ \cos \theta_{\tilde{\mathbf{a}}\tilde{\mathbf{a}}'} + (I \cdot \tilde{\mathbf{c}}) \sin \theta_{\tilde{\mathbf{a}}\tilde{\mathbf{a}}'} \}, \end{aligned} \quad (44)$$

with $\tilde{\mathbf{c}} = \tilde{\mathbf{a}} \times \tilde{\mathbf{a}}' / |\tilde{\mathbf{a}} \times \tilde{\mathbf{a}}'|$. Now suppose the pair $\tilde{\mathbf{b}}\tilde{\mathbf{b}}'$ of vectors starts out being aligned with the pair $\tilde{\mathbf{a}}\tilde{\mathbf{a}}'$, and then gets rotated counterclockwise by angle $\theta_{\tilde{\mathbf{a}}\tilde{\mathbf{b}}}$ relative to the pair $\tilde{\mathbf{a}}\tilde{\mathbf{a}}'$, about the $\tilde{\mathbf{c}}$ -axis. Such a rotation within S^3 can be described *intrinsically*, as

$$\begin{aligned} (+I \cdot \tilde{\mathbf{b}})(+\boldsymbol{\mu} \cdot \tilde{\mathbf{b}}') &= \mathcal{R}_{\tilde{\mathbf{a}}\tilde{\mathbf{b}}} \{ (+I \cdot \tilde{\mathbf{a}})(+\boldsymbol{\mu} \cdot \tilde{\mathbf{a}}') \} \\ &= -\lambda \{ \cos(\theta_{\tilde{\mathbf{a}}\tilde{\mathbf{b}}} + \theta_{\tilde{\mathbf{a}}\tilde{\mathbf{a}}'}) + (I \cdot \tilde{\mathbf{c}}) \sin(\theta_{\tilde{\mathbf{a}}\tilde{\mathbf{b}}} + \theta_{\tilde{\mathbf{a}}\tilde{\mathbf{a}}'}) \}, \end{aligned} \quad (45)$$

where $\mathcal{R}_{\tilde{\mathbf{a}}\tilde{\mathbf{b}}} = \exp \{ (I \cdot \tilde{\mathbf{c}}) \theta_{\tilde{\mathbf{a}}\tilde{\mathbf{b}}} \} = \tilde{\mathbf{a}}\tilde{\mathbf{b}}$ is the rotor that parallel transports the quaternion $(+I \cdot \tilde{\mathbf{b}})(+\boldsymbol{\mu} \cdot \tilde{\mathbf{b}}')$ from “point $\tilde{\mathbf{a}}$ ” to “point $\tilde{\mathbf{b}}$ ” within the 3-sphere [7][8][14]. But using the identity

$$\cos(\theta_{\tilde{\mathbf{a}}\tilde{\mathbf{b}}} + \theta_{\tilde{\mathbf{a}}\tilde{\mathbf{a}}'}) = \cos \theta_{\tilde{\mathbf{a}}\tilde{\mathbf{b}}} \cos \theta_{\tilde{\mathbf{a}}\tilde{\mathbf{a}}'} - \sin \theta_{\tilde{\mathbf{a}}\tilde{\mathbf{b}}} \sin \theta_{\tilde{\mathbf{a}}\tilde{\mathbf{a}}'},$$

and noting that $I \cdot \tilde{\mathbf{c}}$ in the limit $\tilde{\mathbf{a}}' \rightarrow \tilde{\mathbf{a}}$ reduces to a null bivector, it is easy to deduce from this rotation that

$$\begin{aligned} \mathcal{B}(\beta, \boldsymbol{\mu}) &= \lim_{\tilde{\mathbf{b}}' \rightarrow \tilde{\mathbf{b}}} [(+I \cdot \tilde{\mathbf{b}})(+\boldsymbol{\mu} \cdot \tilde{\mathbf{b}}')] = -\lambda \\ &\longrightarrow \lim_{\tilde{\mathbf{a}}' \rightarrow \tilde{\mathbf{a}}} [\mathcal{R}_{\tilde{\mathbf{a}}\tilde{\mathbf{b}}} \{ (+I \cdot \tilde{\mathbf{a}})(+\boldsymbol{\mu} \cdot \tilde{\mathbf{a}}') \}] = -\lambda \cos \theta_{\tilde{\mathbf{a}}\tilde{\mathbf{b}}} = -\lambda \cos 2(\alpha - \beta). \end{aligned} \quad (46)$$

This is a probabilistic prediction that says that the relative frequency of occurring the result $\mathcal{B}(\beta, \boldsymbol{\mu}) = -\lambda$ will vary with the cosine of the angle $\theta_{\tilde{\mathbf{a}}\tilde{\mathbf{b}}}$. In particular, as $\tilde{\mathbf{b}} = \tilde{\mathbf{a}} \rightarrow -\tilde{\mathbf{a}}$, the result $\mathcal{B}(\beta, \boldsymbol{\mu})$ will change from $-\lambda$ to $+\lambda$, and consequently the value of the product $\mathcal{A}(\alpha, \boldsymbol{\mu}) \mathcal{B}(\beta, \boldsymbol{\mu})$ will change from -1 to $+1$, just as we discussed above. This, then, is a manifestation of the “twist” in the Hopf fibration of the 3-sphere mentioned above [14][15][16]. By contrast, had we modeled the physical space as \mathbb{R}^3 instead of S^3 , the same result would have appeared as a nonlocal effect.

In the previous paragraphs we investigated the cases $\tilde{\mathbf{b}} \rightarrow \tilde{\mathbf{a}}$ and $\tilde{\mathbf{b}} \rightarrow -\tilde{\mathbf{a}}$ in some detail. For the general case $\tilde{\mathbf{b}} \neq \tilde{\mathbf{a}}$ it is convenient to view the analyzers $(-I \cdot \tilde{\mathbf{a}})$ and $(+I \cdot \tilde{\mathbf{b}})$ as two different scales of measurements, as we have done above. Accordingly, we may think of polarization $(+\boldsymbol{\mu} \cdot \tilde{\mathbf{a}})$ observed by Alice as scaled by $(-I \cdot \tilde{\mathbf{a}})$ to obtain the raw score $\mathcal{A}(\alpha, \boldsymbol{\mu}) = \pm 1$, and likewise polarization $(+\boldsymbol{\mu} \cdot \tilde{\mathbf{b}})$ observed by Bob as scaled by $(+I \cdot \tilde{\mathbf{b}})$ to obtain the raw score $\mathcal{B}(\beta, \boldsymbol{\mu}) = \pm 1$. But since $(-I \cdot \tilde{\mathbf{a}})$ and $(+I \cdot \tilde{\mathbf{b}})$ are two *different* scales measuring the same orientation of S^3 , the correct correlation between the raw scores can only be determined by comparing the corresponding standard scores, as we have done in Eqs. (30) to (41). Alternatively (but equivalently), the correct correlation between the raw scores can be determined by calculating their covariance divided by the product of their standard deviations, as explained in Eq. (21), with the product of the standard deviations in our case being $(-I \cdot \tilde{\mathbf{a}})(+I \cdot \tilde{\mathbf{b}}) = \tilde{\mathbf{a}}\tilde{\mathbf{b}}$ (see also [18]). This product however is precisely the rotor that quantifies the twist in the fibration of S^3 , as in Eq. (46) above (see also Ref. [14]). Its value varies from $+1$ for $\tilde{\mathbf{b}} = \tilde{\mathbf{a}}$ to -1 for $\tilde{\mathbf{b}} = -\tilde{\mathbf{a}}$ and back, producing the correct combination of probabilities for detection of polarizations $++$, $--$, $+-$, and $-+$ observed in the experiments. Consequently, in accordance with our results (33) and (46), when the raw scores $\mathcal{A} = \pm 1$ and $\mathcal{B} = \pm 1$ are compared in practice by coincidence counts [4][5], the normalized expectation value of their product will inevitably yield

$$\begin{aligned} \mathcal{E}(\alpha, \beta) &= \frac{[C_{++}(\alpha, \beta) + C_{--}(\alpha, \beta) - C_{+-}(\alpha, \beta) - C_{-+}(\alpha, \beta)]}{[C_{++}(\alpha, \beta) + C_{--}(\alpha, \beta) + C_{+-}(\alpha, \beta) + C_{-+}(\alpha, \beta)]} \\ &= \lim_{n \gg 1} \left[\frac{1}{n} \sum_{i=1}^n \mathcal{A}(\alpha, \boldsymbol{\mu}^i) \mathcal{B}(\beta, \boldsymbol{\mu}^i) \right] \\ &= -\cos 2(\alpha - \beta), \end{aligned} \quad (47)$$

where $C_{+-}(\alpha, \beta)$ etc. represent the number of joint occurrences of detections $+1$ along \mathbf{a} and -1 along \mathbf{b} etc.

Now the important feature of the experiments performed in Orsay [4] and Innsbruck [5] was that the measurement settings at the remote polarizers were changed *during* the flight of the two particles, thus removing any possibility of communication between the two ends. The Innsbruck team precluded the possibility of communication between the polarizers by using random and ultrafast switching of the orientations of the polarizers. On each side of the experiment a local computer registered the polarizer orientation and the result of each measurement, with timing monitored by an atomic clock, and the data was gathered and compared for correlation only after the end of each run. These achievements prompted the principal investigator of the Orsay experiment to make the following comment:

I suggest we take the point of view of an external observer, who collects the data from the two distant stations at the end of the experiment, and compares the two series of results. This is what the Innsbruck team has done. Looking at the data a posteriori, they found that the correlation immediately changed as soon as one of the polarizers was switched, without any delay allowing for signal propagation ... [17].

But given the explanation above this immediate change in correlation should be no more puzzling than the sudden change in the color of Dr. Bertlmann's sock discussed by Bell [19]. In fact, in our opinion classical men like Hamilton, Grassmann, Clifford, and Hopf would not have been puzzled by this change at all. They would have explained the results observed in the experiments in exactly the same local terms as we have explained them here; namely, in terms of the algebra of rotations in the physical space S^3 , otherwise known as Clifford algebra [7][8][16][18].

Acknowledgments

I wish to thank Abner Shimony for suggesting that I should explain the results of Orsay and Innsbruck experiments. I also wish to thank the Foundational Questions Institute (FQXi) for supporting this work through a Mini-Grant, and Heine Rasmussen for inspiring the appendix.

APPENDIX: ERROR PROPAGATION WITHIN A PARALLELIZED 3-SPHERE

In this appendix we spell out the statistical basis of the results (25) and (26) in more detail. To this end (recalling that a unit 3-sphere is homeomorphic to a set of unit quaternions), consider a probability density function $P(\mathbf{q}) : S^3 \rightarrow [0, 1]$ of random quaternions over S^3 :

$$P(\mathbf{q}) = \frac{1}{\sqrt{2\pi \|\sigma(\mathbf{q})\|^2}} \exp \left\{ -\frac{\|\mathbf{q} - m(\mathbf{q})\|^2}{2 \|\sigma(\mathbf{q})\|^2} \right\}. \quad (48)$$

It is a matter of indifference whether the distribution of \mathbf{q} so chosen happens to be normal or not [11]. Here $\mathbf{q} = \mathcal{A} + \mathbf{w} = \text{scalar} + \text{bivector}$ is a generic quaternion within S^3 whose mean value $m(\mathbf{q})$ is defined as

$$m(\mathbf{q}) = \frac{1}{n} \sum_{i=1}^n \mathbf{q}^i \quad (49)$$

and whose standard deviation $\sigma(\mathbf{q})$ (with $\|\mathbf{q}\|^2 = \mathbf{q}^\dagger \mathbf{q}$) is defined as

$$\sigma(\mathbf{q}) = \sqrt{\frac{1}{n} \sum_{i=1}^n \|\mathbf{q}^i - m(\mathbf{q})\|^2}. \quad (50)$$

Needless to say, these definitions remain valid for the limiting cases of the random variable $\mathbf{q} \in S^3$ when $\mathbf{q} = \mathcal{A} \in [-1, +1]$ and $\mathbf{q} = \mathbf{w} \in S^2$. Now let $\mathbf{w} = p \boldsymbol{\mu} \cdot \mathbf{n} \in S^3$ be a random bivector with $\|\boldsymbol{\mu} \cdot \mathbf{n}\|^2 = 1$ and $p \in [0, 1]$, where $\boldsymbol{\mu} = \lambda I$ is the indefinite volume form as before and $\mathbf{n} \in \mathbb{R}^3$ is a unit vector. Then, using equation (50), it is

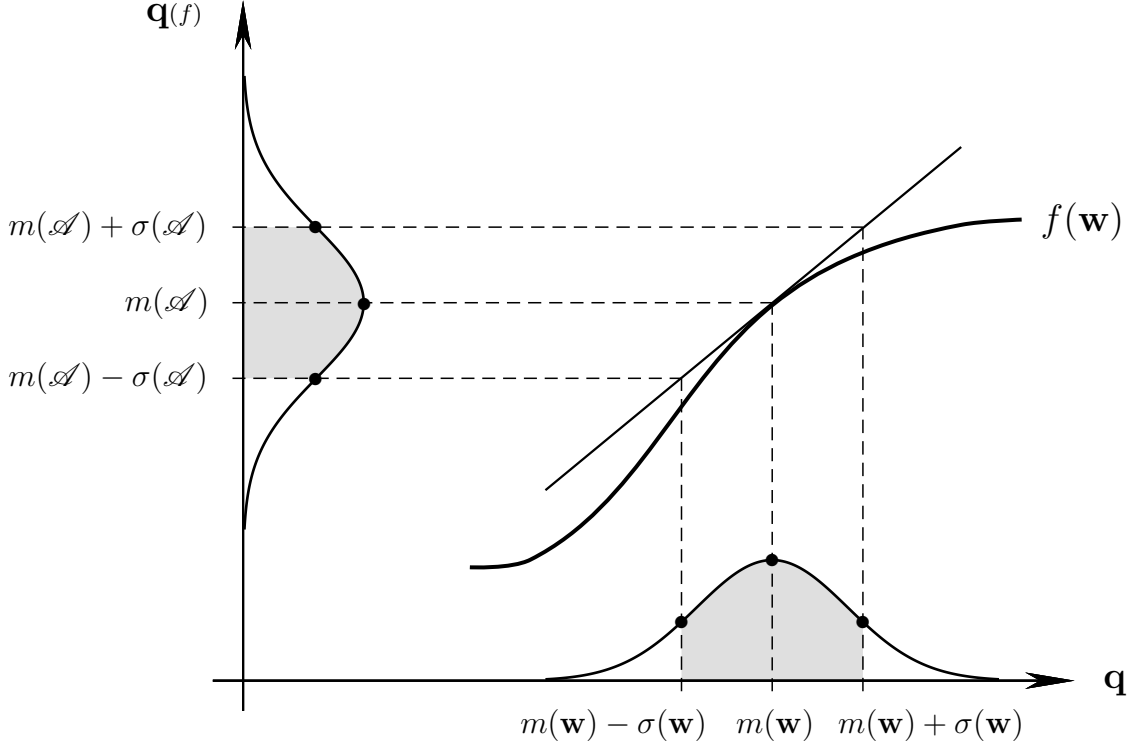


FIG. 1: Propagation of error within a parallelized 3-sphere.

easy to verify that, in general, the mean $m(\mathbf{w})$ of \mathbf{w} and the standard deviation $\sigma(\mathbf{w})$ of \mathbf{w} , respectively, would be a bivector and a scalar:

$$\begin{aligned} m(\mathbf{w}) &= \text{a bivector} \\ \text{and } \sigma(\mathbf{w}) &= \text{a scalar.} \end{aligned} \tag{51}$$

Suppose now we consider another non-random bivector, $\mathbf{v} = I \cdot \mathbf{n}$, and define a scalar number $-1 \leq \mathcal{A} = \mathbf{v} \mathbf{w} \leq +1$, so that $m(\mathcal{A}) \geq 0$. Since \mathbf{v} is a non-random bivector, errors generated within \mathcal{A} by the random process $p\lambda$ stem entirely from the random bivector \mathbf{w} , and propagate linearly. In other words, the standard deviations within the random number \mathcal{A} due to the random process $p\lambda$ is given by

$$\sigma(\mathcal{A}) = \mathbf{v} \sigma(\mathbf{w}). \tag{52}$$

But since $\sigma(\mathbf{w})$ is a scalar, the typical error $\sigma(\mathcal{A})$ generated within \mathcal{A} due to the random process $p\lambda$ is a bivector [18]. The standardized variable (which must be used to compare the raw scores \mathcal{A} with other raw scores \mathcal{B}) is thus also a bivector: $A := \mathcal{A} / \sigma(\mathcal{A}) = \text{scalar} \times \mathbf{w}$.

As straightforward as it is, the above conclusion may seem rather unusual. It is important to recall, however, that in geometric algebra scalars and bivectors are treated on equal footing. They both behave as real-valued c-numbers, albeit of different grades [8]. To appreciate the consistency and naturalness of the above conclusion, let

$$\mathcal{A} = f(\mathbf{w}) = \mathbf{v} \mathbf{w} \tag{53}$$

be a continuous random scalar generated by the geometric product of the two bivectors \mathbf{v} and \mathbf{w} , as before. The natural question then is: How does a typical error in \mathbf{w} governed by the probability density (48)—which can be represented by the 68% probability interval

$$[m(\mathbf{w}) - \sigma(\mathbf{w}), m(\mathbf{w}) + \sigma(\mathbf{w})] \tag{54}$$

as shown in the Fig.(1)—propagate from the random bivector \mathbf{w} to the random scalar \mathcal{A} , through the function $f(\mathbf{w}) = \mathbf{v} \mathbf{w}$? To answer this question we note that the two end points of the interval (54) represent two non-equatorial points, \mathbf{q}^- and \mathbf{q}^+ , of the parallelized 3-sphere, which is a Riemannian manifold. The geometro-algebraic distance between the points \mathbf{q}^- and \mathbf{q}^+ can therefore be defined as

$$d(\mathbf{q}^-, \mathbf{q}^+) = (\mathbf{q}^- - \mathbf{q}^+) \times \text{sign}(\mathbf{q}^- - \mathbf{q}^+). \quad (55)$$

Moreover, from the definition (53) of \mathcal{A} and a first-order Taylor expansion of the function $f(\mathbf{w})$ about the point $\mathbf{w} = m(\mathbf{w})$ we obtain

$$\mathcal{A} = f(m(\mathbf{w})) + \left. \frac{\partial f}{\partial \mathbf{w}} \right|_{\mathbf{w} = m(\mathbf{w})} (\mathbf{w} - m(\mathbf{w})) + \dots \quad (56)$$

Now it is evident that the slope $\partial f / \partial \mathbf{w} = \mathbf{v}$ of this line is a constant. Therefore the mean $m(\mathcal{A})$ and the standard deviation $\sigma(\mathcal{A})$ of the distribution of \mathcal{A} can be obtained by setting $\mathbf{w} = m(\mathbf{w})$ and $\mathbf{w} = \sigma(\mathbf{w})$:

$$\begin{aligned} m(\mathcal{A}) &= f(m(\mathbf{w})) = \mathbf{v} m(\mathbf{w}) = \text{a scalar} \\ \text{and } \sigma(\mathcal{A}) &= \left. \frac{\partial f}{\partial \mathbf{w}} \right|_{\mathbf{w} = \sigma(\mathbf{w})} \sigma(\mathbf{w}) = \mathbf{v} \sigma(\mathbf{w}) = \text{a bivector}. \end{aligned} \quad (57)$$

The probability distribution of \mathcal{A} is thus represented by the interval

$$[m(\mathcal{A}) - \sigma(\mathcal{A}), m(\mathcal{A}) + \sigma(\mathcal{A})]. \quad (58)$$

If we now assume that \mathbf{w} is a unit bivector with a vanishing mean, then we have $m(\mathcal{A}) = 0$ and $\sigma(\mathcal{A}) = \mathbf{v}$, as in equation (23) above.

It is instructive to note here that, geometrically, the propagation of error within S^3 is equivalent to a simple change in perspective:

$$S^3 \ni \underbrace{\overbrace{m(\mathbf{w})}^{\text{bivector}} \pm \overbrace{\sigma(\mathbf{w})}^{\text{scalar}}}_{\text{quaternion}} \xrightarrow{f(\mathbf{w})} \underbrace{\overbrace{m(\mathcal{A})}^{\text{scalar}} \pm \overbrace{\sigma(\mathcal{A})}^{\text{bivector}}}_{\text{quaternion}}. \quad (59)$$

In particular, the probability distribution of \mathcal{A} over S^3 corresponding to (58) is equivalent to that of \mathbf{w} over S^3 corresponding to (54).

-
- [1] A. Einstein, B. Podolsky, and N. Rosen, Phys. Rev. **47**, 777 (1935).
 - [2] J. S. Bell, Physics **1**, 195 (1964).
 - [3] A. Peres, *Quantum Theory : Concepts and Methods* (Kluwer Academic, Dordrecht, 1993), pp. 155.
 - [4] A. Aspect, P. Grangier, and G. Roger, Phys. Rev. Lett. **49**, 91 (1982); See also A. Aspect, J. Dalibard, and G. Roger, Phys. Rev. Lett. **49**, 1804 (1982).
 - [5] G. Weihs *et al.*, Phys. Rev. Lett. **81**, 5039 (1998).
 - [6] W. F. Eberlein, Am. Math. Monthly **69**, 587 (1962); See also W. F. Eberlein, Am. Math. Monthly, **70**, 952 (1963).
 - [7] D. Hestenes, *New Foundations for Classical Mechanics*, Second Edition (Kluwer Academic, Dordrecht, 1999); C. Doran and A. Lasenby, *Geometric Algebra for Physicists* (Cambridge University Press, Cambridge, 2003).
 - [8] J. Christian, arXiv:1101.1958 (see also arXiv:1201.0775).
 - [9] M. Nakahara, *Geometry, Topology and Physics* (Adam Hilger, Bristol, 1990).
 - [10] A. Shimony, Brit. J. Phil. Sci. **35**, 25 (1984).
 - [11] J. L. Rodgers and W. A. Nicewander, The American Statistician **42**, 59 (1988).
 - [12] L. Zulli, Am. Math. Monthly **103**, 221 (1996); See also J. Christian, arXiv:0904.4259.
 - [13] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Phys. Rev. Lett. **23**, 880 (1969).
 - [14] T. Eguchi, P. B. Gilkey, and A. J. Hanson, Physics Reports **66**, 213 (1980) [*cf.* page 272, Eq. (4.21)].
 - [15] L. H. Ryder, J. Phys. A **13**, 437 (1980); D. W. Lyons, Mathematical Magazine **76**, 87 (2003).
 - [16] J. Christian, arXiv:0806.3078.
 - [17] A. Aspect, Nature **398**, 189 (1999).
 - [18] J. Christian, arXiv:1103.1879.
 - [19] J. S. Bell, *Speakable and Unsayable in Quantum Mechanics* (Cambridge University Press, 1987), Chapter 16.